id:analytics®

WHITEPAPER

The Long Con: An Analysis of Synthetic Identities

October 2014

Table of Contents

Introduction

Fraudsters continue to find increasingly sophisticated means to perpetrate crimes despite advances in technology and stricter regulations designed to reduce fraudulent behavior. Notably, the advent of digital technology and the anonymity it provides has contributed to the rise of the creation of fictitious credentials known as synthetic identities. Synthetic identity fraud is a significant and growing problem for multiple reasons. A major contributing factor is that social security number (SSN) randomization, introduced in 2011 and intended to provide higher safeguards for the public, has been exploited by fraudsters to help create and use false identities. The availability of identity information for use by fraudsters in generating synthetic identities is also growing with each new, massive data breach. Additionally, the shift to the Europay MaterCard Visa (EMV) standard for point-of-sale transactions in 2015, which requires merchants to accept both magnetic strip and chip and PIN transactions, is expected to force fraudsters to adopt new account fraud strategies.

In the midst of this changing landscape, it is reasonable to assume that synthetic fraud will continue to be a persistent and significant problem. Avivah Litan, Gartner Inc., estimates that "synthetic schemes constitute at least 20% of credit charge-offs and 80% of losses from credit-card fraud".[1] The true danger of synthetic fraud is that, unlike third-party fraud where an entire identity is stolen and used to defraud enterprises and victims, synthetic fraud frequently has no specific consumer victim. The lack of a clear consumer victim often allows a synthetic fraudster to remain undetected for months, only to "bust out" (suddenly use the remainder of a credit line). This long-term "con" or fraud is particularly dangerous because criminals employing this technique for financial gain can often nurture the synthetic identity into generating larger credit limits and larger loss amounts for the lender than the average identity theft scenario.

This white paper explores 1) the definition of synthetic fraudsters and the algorithm used to identify them, 2) the estimated size of the synthetic fraud issue, 3) various regulatory and financial issues that lenders and wireless carriers face when it comes to synthetic fraudsters, and 4) possible solutions for reducing synthetic fraud.

Background

Fraudsters have used synthetic identities for decades. In the past, people with minimal or poor credit history created synthetic identities to receive approval for new credit cards. In some cases, to consolidate financial responsibility, family members shared identity information in an amalgamation of elements that could no longer be traced to a particular person. Despite the fake identities, these cases often would not derive from any malicious intent, and applicants would plan to use the card legitimately to purchase goods. However, today, synthetic identities are often created for malicious intent: to purposefully commit fraud, to launder money, or to support terrorist or other criminal activities.

The risk of synthetic identity fraud has grown as the demand for immediate access to goods and services threatens to strain security processes. Most enterprise systems have become increasingly automated and online enrollment technologies have become ever more popular. Due to the absence of official identification, like driver's licenses or passports, synthetic fraudsters can appear as a "thin file/no hit" applicant to an enterprise's adjudication system. Without a true paper trail leading to a real person, the subsequent drop in credit score is of no consequence and the victim of the crime is the financial institution at risk of fraud or violation of Know Your Customer (KYC) regulations.

[1] Christopher Conkey, *The Borrower Who Never Was* (The Wall Street Journal, 2007)

Defining Synthetic Identities: Distinctive Characteristics of Synthetic Identity Fraud

Synthetic identities are a combination of fabricated credentials where the implied identity is not associated with a real person. Unlike identity theft or manipulation, where the core identity of a person is impersonated or manipulated by a fraudster, a synthetic identity is an artificial identity with no particular person behind it. The fake identity uses false elements that include any combination of fabricated SSN, name, date of birth, address or phone number. The validity of each individual identity element is not definitive, but the combination of identity elements is definitely invalid. For example, the address may be a real, "shippable" address and the SSN may be real and valid, but the SSN, name and date of birth combination does not match with any real person.

Synthetic identities are fairly easy to create because fabricating data elements can be done indiscriminately. Even an SSN, which is supposed to be a unique identifier for a single person, can be easily fabricated. Blatantly invalid SSNs (e.g. 123-00-0000) would most likely be flagged by a financial institution due to the obvious invalidity of some of the data elements; however, potentially legitimate SSNs are easier to fabricate in the post-randomization era. Because a fabricated SSN may be a previously issued SSN to an unknown person, the resulting fraud may appear to be identity theft. However, this method actually indicates synthetic fraud because the fabricated personally identifiable information (PII) does not belong to the SSN, to the fraudster, or to any core identity.

This accidental or intentional use of a valid SSN with fabricated identity elements is one of the more difficult scenarios for lenders and vendors to identify and prevent. Furthermore, lenders and vendors typically rely on the SSN for identity verification and as long as the SSN does not raise any red flags (frivolous, deceased or invalid format/range), the unknown identity may bypass the lender's security system as an applicant with a "thin" credit history. Some synthetic fraudsters will create a shell company to artificially verify fake identity elements. The use of a valid SSN with other fake identity elements to commit synthetic identity fraud is a common way to attempt to bypass a lender's security measures.

Differentiating Synthetic Identity Fraud

The table below summarizes the characteristics of three distinct identity fraud types that are often confused:

	Identity Theft	Identity Manipulation	Synthetic Identity
Who is the victim?	• The person whose identity is being misused • The company providing the fraudulently obtained product or service is also a victim	• The fraudster is the victim • The company providing the fraudulently obtained product or service is also the victim	• No consumer victim • The company providing the fraudulently obtained product or service is the victim
Nature of the misrepresentation	• Typically SSN, name and date of birth belong to the victim, and the address, phone, and/or email are associated with the fraudster	• SSN, date of birth and/or name vary slightly from the fraudster's own, correct information	• SSN, name and date of birth are fabricated or chosen randomly
Signals used to find the fraud	• Unusual activity around the SSN, name, and/or date of birth combination (these belong to the victim), or address/phone (these are usually associated with the fraudster)	• Systematic variations around the fraudster's personally identifiable information (PII) to differentiate from simple typos	• Combined identity elements do not match • Asserted PII, especially SSN, has been seen in relation to other identity elements
Summary	• Victim's core identity	• Fraudster's core identity	• No core identity

Ultimately, the differences between identity theft, identity manipulation, and synthetic identity fraud manifests in how the fraudster obtains or creates the asserted PII, including the SSN, and the resulting indicators of fraud. Identity thieves knowingly steal a victim's core identity and use this legitimate combination of SSN, name and date of birth to apply for a product or service. Identity manipulators make slight variations to their own, or a known, identity to attempt to bypass a lender's verification system. Synthetic fraudsters create an identity using a potentially valid SSN and fabricated PII. Because identity theft has a particular consumer victim, there is a much higher likelihood of the fraud being reported to the financial institution since the extra account will eventually appear in the victim's credit report. With manipulation and synthetics, discovery is likely to occur long after the financial loss has been perpetrated, and will generally not be reported by the consumer.

According to a 2007 article in the Harvard Journal of Law & Technology[2]: "However, as estimates of the prevalence and severity of identity theft suggest, it is plain that the market has thus far failed to address the problem. The rise of synthetic identity theft indicates that financial institutions are not authenticating the identities of credit applicants. Instead, it appears that financial institutions are only authenticating the SSN by comparing it to the date of birth, rather than ensuring that the number is issued to the correct person. This means that lenders are not using all the tools available to them to prevent identity theft — simply matching the name of the applicant to the SSN would in many cases make this type of fraud impossible."

[2]"Identity Theft: Making the Known Unknowns Known", Harvard Journal of Law & Technology; Volume 21, Number 1, Page 116, Fall 2007; Chris Jay Hoofnagle

Examples of Synthetic Fraud and Financial Risk

In August 2006, James J. Rose was indicted for devising a scheme to defraud multiple financial institutions issuing credit cards. He and his partner used a small consumer reporting agency to steal valid SSNs to create synthetic identities. One of his synthetic frauds includes using Hilal Salifual[3]'s real SSN to apply for a credit line under the name Hannah Crabtree[4], a fabricated name. He and his partners obtained over 250 credit cards from 15 different banks, charging roughly $760,000 to these synthetic identities[5].

Other behaviors used by synthetic fraudsters are known as **bust-outs**. Bust-out is the term used to describe the behavior seen when a fraudster increases the spending limit of a credit line by paying off small purchases with the goal of obtaining larger loans. Eventually, the fraudster will "bust-out" the entire credit line, acquiring a large loan with no intention of paying it back.

In February 2013, federal agents arrested 13 people for creating roughly 7,000 fake identities to steal roughly $200 million from multiple lenders. The perpetrators found unused SSNs sourced via Craigslist listings that solicited individuals who were not credit active to willingly offer up their credentials[6]. The fraudsters used "drop addresses" as the mailing addresses for the false identities, and created sham companies that accepted credit card payments. After obtaining the cards, the fraudsters "started making small charges and paying off the cards to raise their credit limits".[7] This enabled the fraudsters to incur a large loan that would not have been possible without first making smaller payments, which led to significantly greater financial loss for the lender. As the risk of synthetic fraud grows, financial institutions should remain wary of the increased risk of "bust out" behavior that can originate from synthetic identities.

Tax Fraud

Recent ID Analytics research with a state tax agency showed that, over the span of three years, roughly 1.4% of the tax return population appeared to be synthetic. These synthetic identities resulted in a total tax refund amount of $20 million. The SSNs asserted by these identities had never been seen in combination with the asserted name and date of birth, leading ID Analytics to believe these identities are synthetic.

[3]Name has been anonymized to protect identity's privacy
[4]Name has been anonymized to protect identity's privacy
[5]Chris Jay Hoofnagle, *Identity Theft: Making the Known Unknowns Known* (Harvard Journal of Law and Technology, 2007)
[6]U.S. Attorney's Office, *Four Plead Guilty in $200 Million International Credit Card Fraud Conspiracy* (FBI, 2013)
[7]U.S. Attorney's Office, *Eighteen People Charged in International $200 Million Credit Card Fraud Scam* (FBI, 2013)

Synthetic Fraud and Criminal Activities

Perhaps equally troubling is the potential of other crimes committed under fabricated identities. As ID Analytics' original research uncovered, synthetic identities are used for criminal activity in both the wireless and financial industries. Because these synthetic identities are not used for financial gain, they are much more difficult to discover since there is no report of a fraud or credit loss.

There have been recent, high-profile cases of terrorist organizations exploiting the use of synthetic identities to serve their ideological purposes. Synthetic identities provide an avenue for terrorists to not only distribute funding, but to also obtain valuable resources, such as cell phones and airplane tickets for individuals intent on more than just financial harm. For example, in May of 2014, Canadian officials discovered that terrorists on "do not fly lists" were using synthetic identities to purchase airline tickets between New York, Toronto and Pakistan.[8] The scheme included the use of fake names to obtain passports for an infamous murder suspect and major drug trafficking groups.[9]

Financial institutions should be especially concerned about synthetic frauds due to Know Your Customer (KYC) compliance policies. The USA PATRIOT Act of 2001 requires all financial service providers to establish anti-money laundering programs. Because most synthetic fraudsters appear as "new-to-credit" or "thin file", lenders are not able to verify the asserted identity as well as they would a "thick file" applicant (e.g., someone with multiple trade lines). The fraudster is allowed to complete the application process because the synthetic identity meets the adjudication criteria for a "thin-file / no-hit" prospect. As long as synthetic fraudsters with criminal intent continue to keep their account(s) in good standing, criminals can engage in money laundering or terrorist funding.

Synthetic Identity: Algorithm for Discovery

Original research of historic data made available in ID Analytics' proprietary ID Network(R) has produced an algorithm that identifies synthetic identities by determining whether or not the particular combination of PII has ever been seen before. The algorithm relies on comparing the asserted PII to that which is stored in the ID Network and looking for mismatches between SSN, name and date-of-birth (DOB) combinations. In order to *confidently* identify a synthetic identity, the core combination of identity elements cannot have been seen before and the SSN has historically been used in combination with different PII than the asserted identity elements. If these conditions are met, the identity is not just an applicant without credit history; the identity is likely fake.

Case Study #1: Synthetic Identity (Confidently Synthetic)

"Charles Smith"[10] applied for a credit card for the first time in 2010. Before this application, "Charles Smith" was never seen before. "Charles Smith" used a valid SSN to apply for a credit line; however, in 2003, that same SSN was used to apply for two retail cards for a completely different identity: Jon Trufante[11]. Jon Trufante and "Charles Smith" have different names, dates of birth, addresses, and phone numbers. In fact, the addresses they used to apply for credit are nearly 100 miles apart.

[8]Rick MacInnes-Rae and Mark Gollom, *Suspected terrorist links to synthetic ID fraud are being 'ignored'* (CBCNews, 2014)
[9]Dave Seglins and John Nicol, *RCMP bust passport fraud scheme tied to Canada's 'most vanted'* (CBCNews, 2014)
[10]Name has been anonymized to protect the identity's privacy.
[11]Name has been anonymized to protect the identity's privacy.

"Charles Smith" fulfills the algorithm for identifying a synthetic identity. He has never been seen before this application and the SSN he is asserting appears to belong to a completely different person. As further evidence of fraud, Jon Trufante again applied for credit from another credit card company three months after "Charles Smith" appeared. This strongly suggests that "Charles Smith" used Jon Trufante's SSN to open a new account with a major lender, but because the new account uses fabricated, unshared, credentials, it is synthetic fraud, not identity theft.

Emerging Market, Immigrant or Possible Synthetic Identity

Emerging Market, Immigrant or **Possible Synthetic**	• SSN never seen before
	• Name and DOB combination never seen before

The algorithm ID Analytics uses to identify synthetic identities is purposefully created to minimize false positives. Synthetic identities have the potential to hide in the "thin file" category due to the lack of credit history surrounding the SSN, name, and date of birth. This population has no previous credit history and has never applied for a phone in the U.S., but they can be completely valid applications. These identities can be emerging market (e.g., "youths"), immigrant, or possible synthetic. Without any historical data to verify the identity elements used in conjunction with the SSN, it is extremely difficult to identify which identities are simply new and which are fabricated. This population of possible synthetic identities is expected to grow in the future due to SSN randomization.[12] Prior to 2011, asserted SSNs from the range of numbers that the Social Security Administration (SSA) never legitimately issued were treated with special scrutiny. After 2011, SSNs that once fell in the SSA's "unissued range" are now legitimately being issued to newborns and immigrants. These applications, while most likely benign, should be scrutinized more closely in order to mitigate the possibility of a synthetic identity making it through the application process.

Case Study #2: Emerging Market or Possible Synthetic Identity

Sarah McMaster[13] applied for a phone at a major wireless carrier in 2010. At the time of the application, Sarah McMaster was under 20 years old. This phone application is likely the first thing Sarah McMaster applied for that required a credit check due to her age. Since 2010, Sarah has continued to apply for multiple credit lines, including a retail card and a major credit card. She has used the same core PII, with some variations in address and phone number.

This identity can likely be categorized as emerging market. The individual is young and had no credit history before applying for her first phone. She also continued to apply for credit using the same SSN, name and DOB combination after the initial phone application. Sarah is most likely a legitimate youth applying for her first phone; however, Sarah may still be a synthetic fraudster due to her lack of credit history. Without historical credit information, it is difficult to conclude for certain that Sarah is a legitimate identity. Until Sarah perpetrates a loss to the wireless carrier, it will remain unclear as to whether she represents a credit risk (credit loss as a legitimate identity) or a fraud risk (potential synthetic identity).

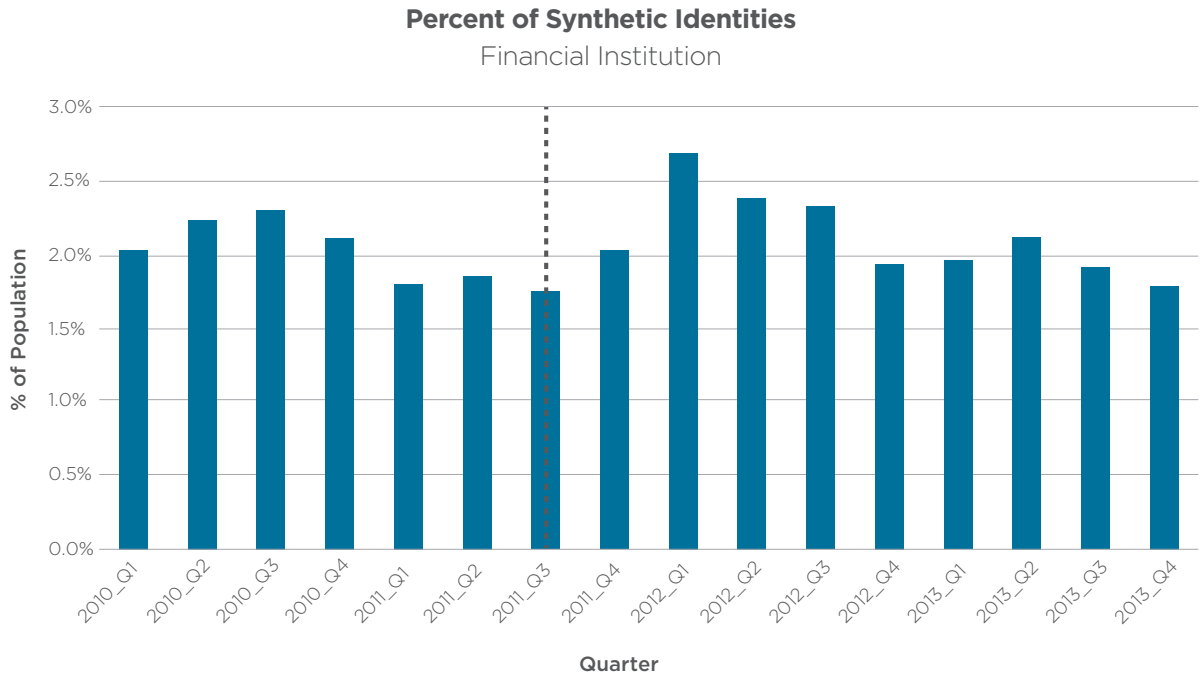
[12]ID Analytics, *Exploring the Impact of SSN Randomization* (ID Analytics, 2014)
[13]Name has been anonymized to protect the identity's privacy.

Magnitude of Volume and Risk

ID Analytics performed multiple studies with respect to data in the ID Network using the above definitions and algorithm to answer two questions: how large is the population of synthetic identities and how risky is this population? Data from both a credit card issuer and a wireless carrier were studied.

Studies performed on new account applications from these two different enterprises indicate that roughly 2% of the consumers from both industries are identified as a synthetic identity, and these consumers are roughly four times more likely to be reported as fraud in both industries.

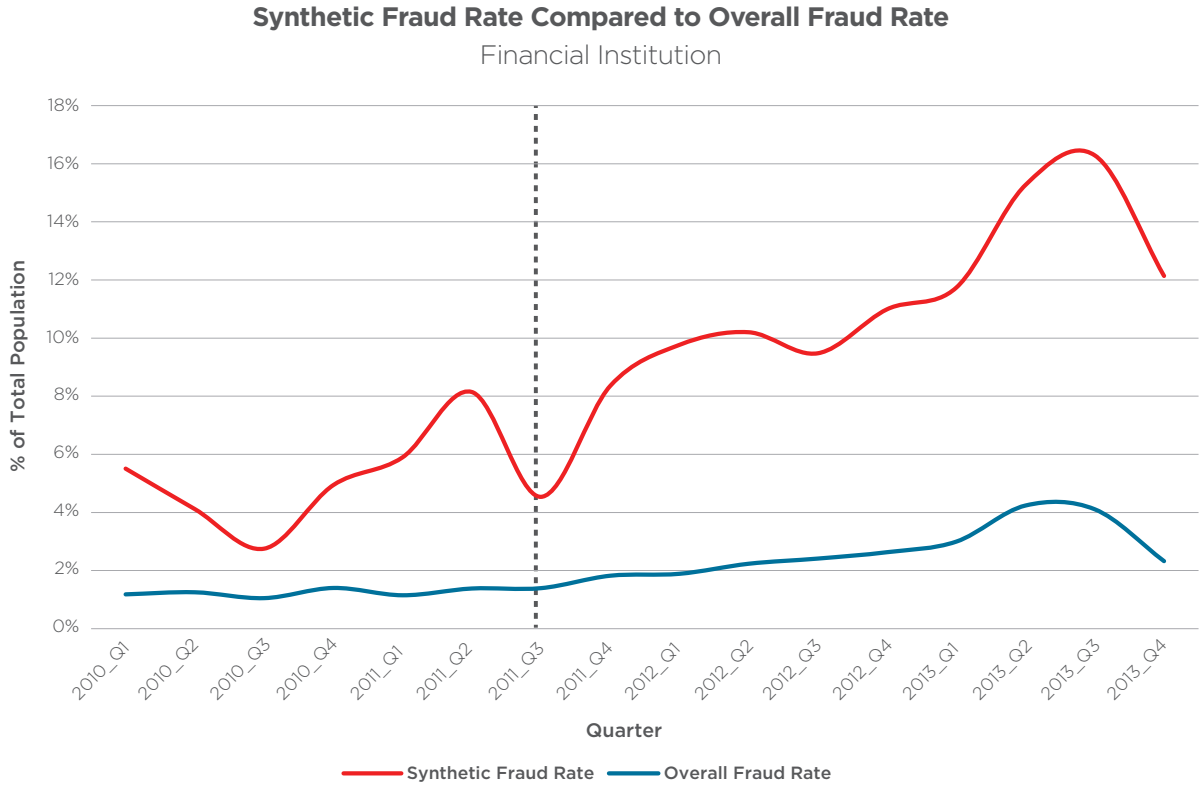
Financial Institutions



The chart above shows the percent of synthetic identities over three years at a credit card issuer. The blue bars represent the percent of total applications identified by ID Analytics as synthetic. The dotted dark gray line approximates the date of SSN randomization, July 2011.

Using ID Analytics' proprietary methodology for categorizing synthetic fraudsters, it is observed that for a credit card issuer, roughly 2% of the total application volume over three years is made up of synthetic identities. The data suggests that over time, the percent of synthetic fraudsters in the total population has decreased slightly. The confounding factor is that there is an increase in the number of SSNs from the previously unissued range due to SSN randomization that may or may not be from synthetic identities.[14]

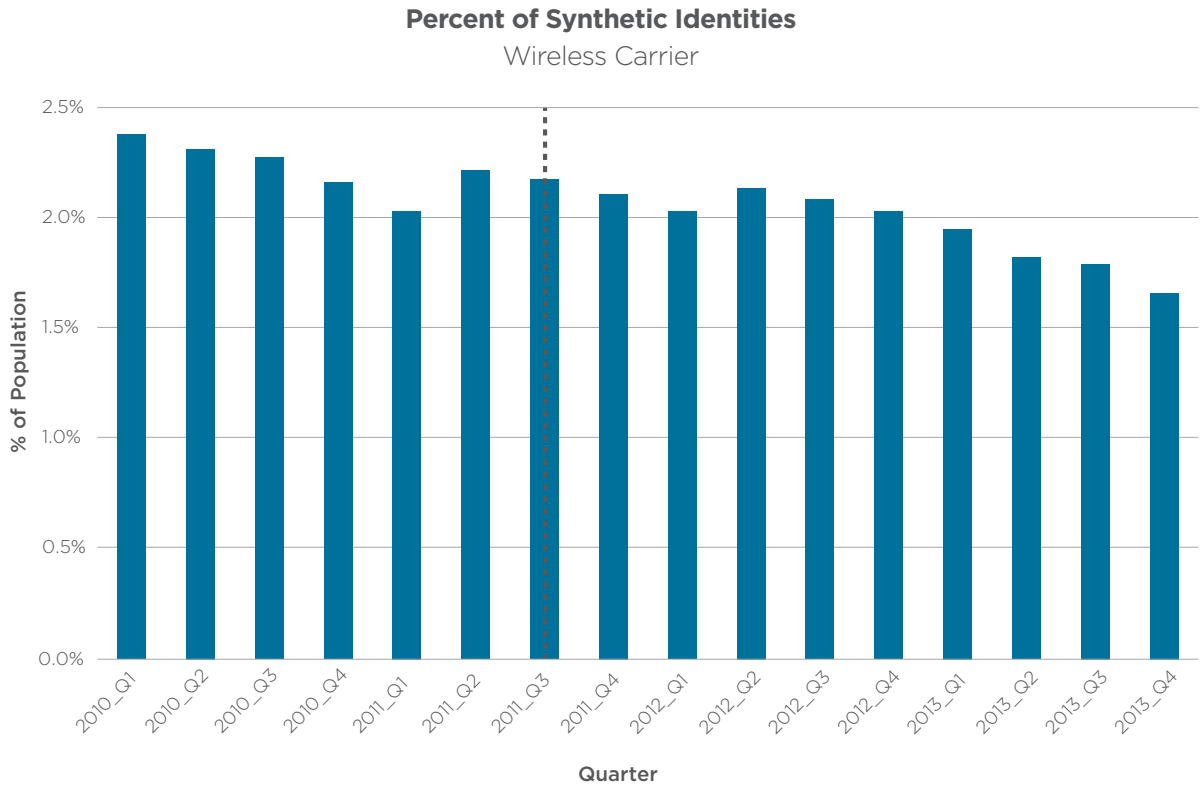
[14]ID Analytics, *Exploring the Impact of SSN Randomization* (ID Analytics, 2014)



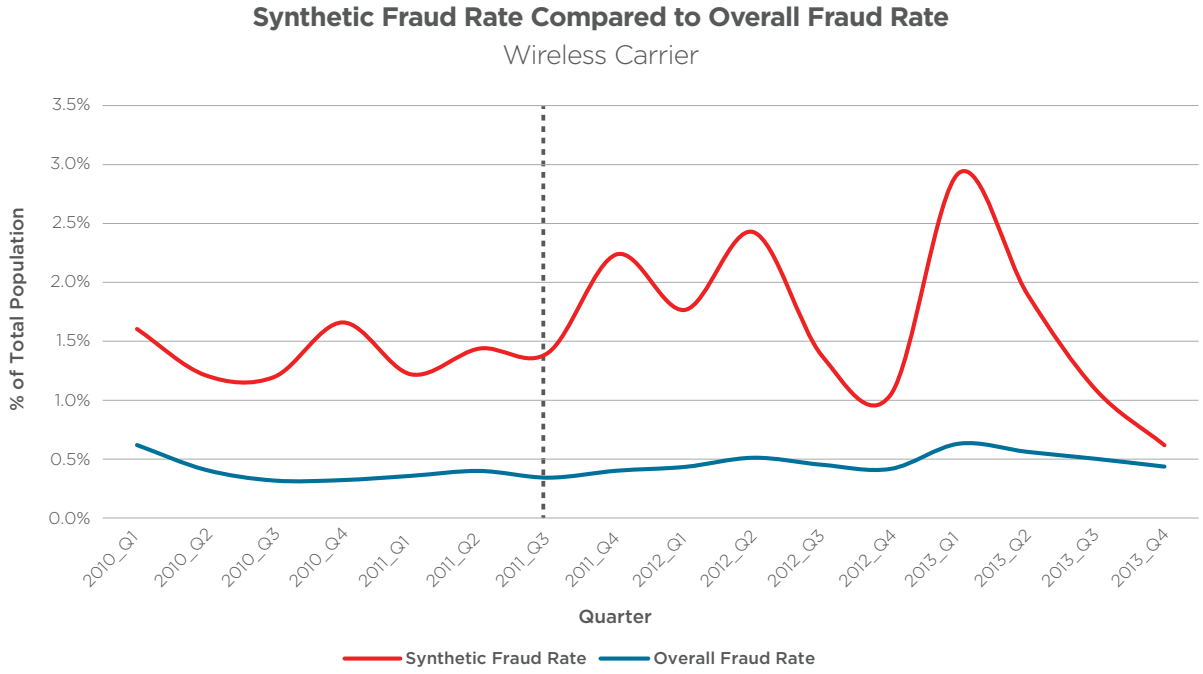
The blue line is the entire set of new-account applications that were discovered to be fraud by lenders (e.g., "tagged as fraud") divided by the entire population of applications. The red line is the set of all tagged frauds for only the subset of applications meeting ID Analytics' criteria for a synthetic identity divided by that entire subset of applications, including those that were not tagged as fraud by the lender.

The chart above shows two fraud rates for two different populations: the synthetic identity population (in red), a subset of the overall population, and the overall population (in blue). The graphic illustrates how much riskier synthetic frauds are than the overall fraud rate (which includes third-party, first-party and synthetic frauds). While the overall population has a fraud rate of roughly 2%, the synthetic identities have an overall fraud rate of about 9%, or over four times riskier. Furthermore, despite a slight decrease in the number of synthetic identities, the synthetic-identity-fraud rate shows signs of increasing over time. This trend coincides with timing of SSN randomization, supporting the supposition that this event increased the risk of synthetic fraud.

Wireless Carriers



The chart above shows the percent of synthetic identities in the total population of a wireless carrier. The data suggests that synthetic identities consist of roughly 1.5 – 2.5% of the total population. Similar to the credit card issuer, the percent of synthetic identities appears to be decreasing over time.



The chart above shows a comparison between the overall fraud rate and the synthetic identity population's fraud rate. Although the wireless carrier's synthetic identity fraud rate has some variability, the rate is consistently above the fraud rate for the total population. The wireless carrier's overall fraud rate is roughly 0.5%, whereas the synthetic identity fraud rate is roughly 2%, about four times riskier than the overall population. Although the trend is not as clear as in the financial institution example, the wireless trend supports the noticeable corresponding rise in the synthetic identity fraud rate following SSN randomization.

It is worth noting that although the synthetic identity population is riskier, there is still a significant portion of the synthetic population that is not flagged as fraud, especially in the wireless carrier's population. This "good" population appears to have every intention to pay. As noted above, there are several possible explanations for these lower-risk applicants. Some applicants may feel as though they must fake their credentials because they have no alternatives; they may be undocumented workers or people escaping bad credit histories. Some may be account holders who are waiting for an opportunity to "bust-out". Others, however, may be using fake identities to maintain accounts that support other illicit activites, but do not necessarily represent a dollar-loss risk to the enterprise. In each case, the synthetic identity appears to be "good" because the person behind the identity makes timely payments on their balances.

Can Synthetics be Recognized?

Synthetic identities can be identified with a high degree of confidence using sufficient historical data and the algorithms that ID Analytics has constructed. Historical application data, especially data on the asserted SSN, is the key to determining if an identity is artificial. The ability to examine all previous name and date of birth combinations associated with an asserted SSN enables identification of synthetic credentials. With a large enough data source like ID Analytics' ID Network, it is possible to confidently determine whether an identity is synthetic or not. Finally, ID Analytics' methodology can categorize identities as likely synthetic identities, possible synthetic identities for further consideration and normal identities.

Conclusion: How Can ID Analytics Help?

With advanced knowledge of the growing threat of synthetic fraud, ID Analytics created the ID Network — a consortium of consumer behavioral data built through the contributions of more than 200 enterprise clients. With more than a trillion data elements from multiple business sectors and information on over three million client-reported attempts at identity fraud, ID Analytics is well situated to identify synthetic fraud. ID Analytics' fraud protection products use historical data to assess the risk of each individual identity element. This is an especially predictive method of flagging synthetic fraud. ID Analytics has historical insight into the vast majority of SSNs and can *confidently* flag identities that appear to be synthetic. ID Analytics can assess the SSN, name and date of birth combinations asserted in the applications and confirm that the same combinations have been used in the past. Using products such as ID Score(R) 9.0 and ID Network Attributes, ID Analytics is well positioned to help solve the synthetic fraud problem.

About ID Analytics, Inc.

ID Analytics is a leader in consumer risk management with patented analytics, proven expertise and real-time insight into consumer behavior. By combining proprietary data from the ID Network®—one of the nation's largest networks of cross-industry consumer behavioral data—with advanced science, ID Analytics provides in-depth visibility into identity risk and creditworthiness. Every day, many of the largest U.S. companies and critical government agencies rely on ID Analytics to make risk-based decisions that enhance revenue, reduce fraud, drive cost savings and protect consumers. ID Analytics is a wholly-owned subsidiary of LifeLock, Inc. Please visit us at www.idanalytics.com.

id:analytics®

www.idanalytics.com